Automated quantitative evaluation of brain MRI may be more accurate for discriminating preterm born adults

Article type Original Article

Abstract

Objective To investigate the structural brain abnormalities and their diagnostic accuracy through qualitative and quantitative analysis in term born and very preterm birth or with very low birth-weight (VP/VLBW) adults.

Methods: We analyzed 3T MRIs acquired 2011-2013 from 67 adults (27 term born controls, mean age 26.4 years, 8 females; 40 VP/VLBWs, mean age 26.6 years, 16 females). We compared automatic segmentations of white matter, deep grey matter and cortical grey matter, manual corpus callosum measurements and visual ratings of the ventricles and white matter with t-tests, logistic regression and receiver operator characteristic (ROC) curves.

Results: Automatic segmentation correctly classified 84% of cases; visual ratings correctly classified 63%. Quantitative volumetry based on automatic segmentation revealed higher ventricular volume, lower posterior corpus callosum and deep grey matter volumes in VP/VLBW subjects compared to controls ($P<0.01$). Visual rating and manual measurement revealed a thinner corpus callosum in VP/VLBW adults ($P=0.04$), deformed lateral ventricles ($P=0.03$) and tendency towards more "dirty" white matter ($P=0.06$). Automatic/manual measures combined with visual ratings correctly classified 87% of cases. Stepwise logistic regression identified three independent features that correctly classify 81% of cases: ventricular volume, deep grey matter volume and white matter aspect.

Conclusion: Enlarged and deformed lateral ventricles, thinner corpus callosum and "dirty" white matter are prevalent in preterm born adults. Their visual evaluation has low diagnostic accuracy. Automatic volume quantification is more accurate but time consuming. It may be useful to ask for prematurity before initiating further diagnostics in subjects with these alterations.

Keywords: preterm birth; low birth weight; adults; Magnetic Resonance Imaging

Abbreviations:

VP/VLBW very preterm birth or very low birth weight

FLAIR fluid attenuated inversion recovery

T1- or T2-w T1 or T2 weighted

PACS picture archiving and communication system

ROC receiver operator characteristic

AUC area under the curve

Keypoints:

- Our study confirms prior reports showing that structural brain abnormalities related to preterm birth persist into adulthood
- In the clinical practice, if large and deformed lateral ventricles, small and thin corpus callosum, and “dirty” white matter are visible on MRI, ask for prematurity before considering other diagnoses
- Although prevalent, visual findings have low accuracy; adding automatic segmentation of lateral ventricles and deep grey matter nuclei improves the diagnostic accuracy

Introduction

For nearly three decades, the rate of preterm birth (before 37 weeks of gestation) and/or low birth weight (below 2500g) steadily increased worldwide [1–3]. Currently, 10% of all live births in developed countries are preterm [1, 4, 5]. Concomitantly, the survival rates for very early preterm births have grown and radiologists are increasingly exposed to adult survivors of preterm birth [6], particularly since especially survivors of very preterm birth or very low birth weight (VP/VLBW, <32 weeks of gestation and/or birth weight <1500g) are predisposed to long-term neurodevelopmental disabilities that persist into adulthood [6–12]. Therefore, knowledge of brain structural changes due to preterm birth is crucial to avoid misdiagnoses.

The immature brain is susceptible to the consequences of preterm birth and abnormal brain development may contribute to neurodevelopmental symptoms that manifest throughout life after preterm birth [13]. MRI detects structural and functional brain alterations related to preterm birth [6, 8, 14–16]. In children, conventional MRI uncovers abnormal signal intensity in white matter, particularly in periventricular areas and along the visual pathways, undulating ventricular borders, ventriculomegaly, cerebral atrophy, thinning of the corpus callosum, and delayed myelination [15, 17, 18]. Special MRI methods such as relaxometry, water fraction or diffusion measurements relate the white matter lesions to axonal or oligodendrocyte injury and abnormal structural connectivity [19–21] while functional MRI relates these structural findings to impaired functional connectivity [22]. Volpe [23] coined the sum of these white and grey matter abnormalities as “the encephalopathy of prematurity” and has identified it as the principal determinant of neurodevelopmental outcome.

Brain abnormalities are well documented in newborns, infants and children and persist into the adult life [12]. VLBW adults have a higher incidence of neurosensory deficits, a greater burden of illness, lower IQ scores, and poorer educational achievement than their peers born with normal birth weight [12, 24–26] and lower quality of life [27]. However, few studies describe the adult preterm brain from a neuroradiological perspective [28–31]. We hypothesized that the brain abnormalities still present in adulthood would form a specific MRI pattern related to premature birth. Therefore, we aimed to investigate the structural brain abnormalities and the diagnostic accuracy of their qualitative and quantitative analysis in term born and VP/VLBW adults.

Methods

Subjects

This study is part of the Bavarian Longitudinal Study (BLS), a prospective, geographically defined whole-population study of neonatal at risk children, who were followed-up from birth into early

adulthood [32–36]. The assessment in adulthood included brain MRI – where this was feasible - performed between Jan 2011 and Dec 2013 at two sites and this is a cross-sectional analysis of the MRI data of all the 67 subjects at one of the sites: 40 VP/VLBW and 27 term born controls (Table 1). The allocation into one of these groups was done based on information collected at birth. By design, birth weight and gestational age were significantly lower in the VP/VLBW group. IQ (Full-Scale IQ) was assessed with the German version of the Wechsler Adult Intelligence Scale (WAIS III) in adulthood [37].

The study was approved by the local ethics committees and all participants gave written informed consent.

Image acquisition

MRI assessments were carried out on a 3T scanner (Philips Achieva or Ingenia, Philips Healthcare) with an 8-channel head coil. Whole brain, high-resolution isotropic T1-weighted (T1-w), 3D fluid-attenuation inversion recovery (FLAIR) and 3D T2-weighted (T2-w) images were acquired (Supplementary Table 1). For both, FLAIR and T2-w sequences, we used a fast spin echo acquisition, based on the variable refocusing flip angle technique (sweep technique) [38].

Data analysis

We evaluated the MR images *quantitatively* through automatic segmentation, manual measurements and *qualitative* ratings of white matter and lateral ventricles aspect. All evaluations were done blinded to group membership (subject's birth status) (Figure 1).

For the *quantitative analysis*, a trained neuroscientist (AJ with 14 years of experience) automatically segmented and analyzed the volumes of lateral ventricles, white matter and grey matter (cortical and deep), and midline corpus callosum (anterior and posterior halves), all normalized for individual head size. This was done using FSL [39] (FMRIB Software Library v5.0, https://fsl.fmrib.ox.ac.uk/fsl) and its tools SIENAX and FIRST. The main steps involved extracting brain and skull images from the single whole-head 3D T1-weighted data, affine-registration to Montral Neurological Institute (MNI) space (using the skull image to determine the registration scaling), tissue-type segmentation with partial volume estimation, masking with standard space masks (for lateral ventricles and corpus callosum) and finally, extraction of normalized segmented volumes.

Manual measurements and qualitative analysis were done on the local PACS, on CE certified diagnostic monitors. Blinded to subject's birth status, a board certified neuroradiologist (EH with 16 years of experience) and a resident radiologist (MM with 5 years of experience) consensually evaluated the T1w, FLAIR and T2w 3D images — with regard to the width of the corpus callosum, the aspect of the lateral ventricles (normal or deformed/enlarged) and presence/absence of white matter lesions or abnormalities (Figure 1). Hereby they manually measured the widths of splenium, isthmus, body and genu of the corpus callosum. They also visually noted the presence of deformed lateral ventricles with billowed or bulged lateral walls of the trigone and posterior horn of the lateral ventricle as well as the presence of white matter lesions or of regions with "dirty"-appearing white matter defined as diffuse or patchy moderate T2-hyperintensity that is between lesions and normal-appearing white matter, according to Ge et al [40].

To compute inter-rater agreement, a third rater (VK, a board-certified neuroradiologist with 8 years of experience) performed a separate additional reading of the 3D T1w, T2w and FLAIR images from all the subjects.

Model selection and specification. As the dependent variable was binary (preterm birth) the appropriate generalized linear model was a logistic regression. The quantitative and qualitative variables resulted from automated segmentation, manual measurements and qualitative ratings were used as explanatory variables in several prediction models. A detailed description of all the models is presented in Supplementary Table 2. As the goodness of fit should not be measured in any form of r-square ([41], page 425) the area under the curve (AUC) was chosen as more appropriate for model comparison. As the explanatory variables might be correlated, a model has to be found which yields an acceptable goodness of fit by avoiding collinearity; as collinearity would disguise potential relevant factors as "not statistically significant".

Model sensitivity. The final model comprised the smallest set of explanatory variables yielding a goodness of fit that is not statistically significantly different from the goodness of fit of the full model and insofar fulfilled the "law of parsimony" (Occam's razor).

Statistics

To test for differences between VP/VLBW adults and controls we used independent samples t-tests (continuous variables) and Fisher's exact test (categorical variables). Except for the volume of the lateral ventricles we could not find any statistically significant deviation from a normal distribution (Shapiro-Wilk tests) in any of the other variables involved in the modeling. In the case of volume of the lateral ventricles, the Mann-Whitney U test delivered similar results as the

t-test. We evaluated the inter-rater agreement by means of Krippendorff's alpha coefficient. We analyzed diagnostic accuracy with logistic regression and receiver operator characteristic (ROC) and report performance at an optimal cut-point of 0.5 to maximize simultaneously both sensitivity and specificity. The models' accuracy and data robustness were validated via bootstrapping [42] with 10000 resamples. We used R Statistics (version 3.3.2 for Mac OS X, R Core Team 2016 www.R-project.org) and STATA software (StataIC 14, StataCorp 2017). P-values are presented as uncorrected and corrected for multiple comparisons with the false discovery rate method [43]. For all analyses, we set statistical significance at P-value < 0.05 and we report means and range if not otherwise specified.

Results

Figure 2 depicts the results of the visual rating, of manual measurements and of automatic segmentation. Exemplary images of lateral ventricles from all VP/VLBW and control participants are presented in Figure 3.

Visual evaluation showed that compared with controls, VP/VLBW subjects had significantly larger and deformed lateral ventricles (65% of VP/VLBW and 33% of controls, odds ratio=3.31, P=0.036) with billowed or bulged lateral walls of the posterior lateral ventricle (posterior cella media, trigonum and occipital horn) towards the white matter and tended to have more white matter alterations (52% of VP/VLBW and 26% of controls, odds ratio=3.08, P=0.06). Manual measurements revealed a thinner isthmus of the corpus callosum in VP/VLBW subjects (P=0.048) After correcting for multiple comparisons, manual measurements and visual evaluations were no longer significant.

The results of the inter-rater analysis revealed a poor agreement for the visual evaluation (49% percent agreement, Krippendorff's alpha -0.018, P-value 0.881 on rating the lateral ventricles as deformed and a 63% percent agreement, Krippendorff's alpha 0.214, P-value 0.085 on rating the white matter aspect as altered). However, we found an excellent agreement for the manual measurements (99%, 97%, 96% respectively 97% percent agreement, Krippendorff's alpha 0.832, 0.467, 0.434 respectively 0.672 and all P-values<0.001 for the width of the genu, anterior part, posterior part (isthmus) respectively splenium of the corpus callosum).

All the preterm born subjects (n=40) and two term born subjects (n=2) underwent cranial ultrasound at birth for detection of perinatal brain injury. Of these, four subjects (all preterm)

had signs of intracranial hemorrhage on the ultrasound images, two of grade III, one of grade II and one of grade I. The two subjects with grade III intracranial hemorrhage showed the following associated alterations on the adulthood MRI: T2-hyperintense lesions with marginal gliosis right periventricular and consecutive compensatory enlargement of the right lateral ventricle to this defect in one subject, and prominent enlargement of the lateral ventricles and overall thinned periventricular white matter in the other subject. Adulthood brain imaging of the other two preterm born subjects, with grade I and II intracranial hemorrhage, did not indicate any associated structural or ventricular alterations.

In adulthood, none of the subjects in our cohort had any neurological symptoms of CSF circulation disturbance. The neuro-cognitive performance of the VP/VLBW individuals in adulthood was the subject of another work [25] but we can confirm in our sample that VP/VLBW had significantly lower IQ scores compared to term born controls (P=0.002, Table 1).

Automatic segmentation (normalized for head size) revealed higher lateral ventricles volume (P=0.003), lower posterior calossal volume (P=0.001) and lower deep grey matter volume (P=0.001) in the VP/VLBW group. White matter and cortical grey matter volumes or size of anterior corpus callosum did not differ between VP/VLBW and controls (Table 2). Absolute, not normalized brain volumes were lower in VP/VLBW than in controls (1181 cm^3 in VP/VLBW versus 1280 cm^3 in term born controls, P=0.0007).

A logistic regression model including all variables could correctly classify 87% of cases with 88% sensitivity and 85% specificity. Models that included only automatic segmentation or automatic segmentation and manual measurements performed similarly well as the full model (P=0.32 and P=0.97 respectively, Table 3), with the automatic segmentation achieving the second best performance from all models with an accuracy of 84%. Visual inspection alone had the lowest classification accuracy, 63% and a specificity of 41%. Based on ROC comparison, models that included only neuroradiological parameters (visual ratings and/or manual measurements) were significantly weaker ($P < 0.05$, Table 3).

The variance inflation factor, an index that measures severity of the collinearity (linear relationship) among explanatory variables, showed excessive collinearity that can complicate or prevent the identification of an optimal set of explanatory variables for a statistical model. Including only a subset of representative variables to reduce collinearity and performing a forward stepwise multivariate logistic regression resulted in a model with three independent

variables - volume of deep grey matter, ventricular volume and white matter aspect - that correctly classified 81% of cases with 85% sensitivity and 74% specificity. However this model, was significantly weaker than the full model (P=0.02, Table 3).

Results from the bootstrap validation confirmed the robustness of all our logistic regression models (Table 3). Details of all the logistic regression models analyzed and the variables in the models, including the odds ratio are presented in Supplementary Table 2.

The male proportion in both groups (control: 8 female of 27 = 30%; preterm: 16 male of 40 =40%) showed no statistically significant difference (Fisher's exact test; P=0.4438). A model including sex as a confounder was statistically not significantly different from the model without sex (Likelihood ratio test P=0.1408).

Discussion

This study describes and quantifies the main brain abnormalities in VP/VLBW born adults compared to term born controls focusing on the diagnostic accuracy of the qualitative and quantitative analysis with regard to these brain abnormalities.

More than half of VP/VLBW subjects showed a deformed aspect of the lateral ventricles and alterations of white matter: The shape of the lateral wall of the dorsal lateral ventricles was deformed and enlarged, the white matter appeared 'dirty' (with intermediate signal intensity between lesions and normal-appearing white matter). A similar pattern was previously found in preterm born newborns, infants, children [23, 44], adolescents [45, 46] and adults [31, 34].
Apart for the qualitative alterations apparent on visual inspection, in the VP/VLBW subjects we measured significantly larger lateral ventricles, thinner midline posterior corpus callosum and observed more white matter alterations. Although these findings are prevalent in the VP/VLBW population, the accuracy of the visual inspection alone was only 63% in our study and the addition of manual measurements improved the accuracy to 73%.

Although the inter-rater agreement was excellent for manual measurements, the categorical variables showed a poor inter-rater agreement. Together with the poor accuracy, these results

suggest that visual assessment is inaccurate to reliably predict premature birth in young adults. The absence of a specific preterm MRI pattern was previously suggested [47], when comparing extremely preterm born children and very preterm 19-year olds with age matched controls. Like the present study, they found ventricular dilatation, white matter affection and thinning of the corpus callosum in the preterm born groups but all three groups shared the same morphological pathology, albeit with higher frequencies in premature cohorts [47].

Although statistically not reliable enough to diagnose prematurity in young adults, the knowledge of some morphological features in the brains of adults with preterm birth status could save the radiologist from wrong diagnostic conclusions. Subjects with a thin corpus callosum, "dirty" white matter and deformed lateral wall of dorsal lateral ventricle should be asked for prematurity. Without this knowledge, the bulged ventricles could lead to the false diagnosis of CSF circulation disturbance, which, together with a small corpus callosum could lead the radiologist to consider a malformation. None of the subjects in our cohort had any neurological symptoms of CSF circulation disturbance. However, compensated chronic hydrocephalus, which may be caused by perinatal intraventricular hemorrhage in premature newborns (4/40 in our cohort), cannot be ruled out. However, the prematurity-related ventricle deformation described here differs from ventricle enlargement in conditions with hydrocephalus. In the former, damage of parietal white matter near the lateral ventricles causes an ex vacuo deformation of the lateral ventricle wall in this region. In contrast, increased inner wall tension in hydrocephalus often leads to a ballooning of the ventricle horns instead of the lateral walls, and moreover the horns are often surrounded by T2-hyperintense halos and the third ventricles are ballooned, which was not the case in our subjects investigated here.

Similarly, the "dirty" white matter can lead to the false diagnosis of a white matter disease, such as MS or cerebral small vessel disease. If such findings are present in an adult, the simple question of preterm delivery/low birth weight could avoid misdiagnosis and its consequences.

Other signs suggestive of preterm birth have been described in the literature. Some authors found that preterm born infants have a higher prevalence of head shape abnormalities such as elongated head shape (dolichocephaly) [48]. Additionally, regional biometric differences reflecting impaired cerebellar size or deviating head diameters were reported in children [49] and adults [28] born preterm and may be associated with cognitive and motor outcomes [49]. Of these, our study confirmed the lower brain size of VP/VLBW, which was the reason why we normalized all brain volumes analyzed in this study.

The automatic segmentation with its subsequent analysis of normalized volumes, had a better accuracy (84%) than the visual evaluation and confirmed that VP/VLBW adults have significantly higher volumes of lateral ventricles and smaller volumes of posterior corpus callosum and deep grey matter. However, the use of automatic segmentation is time consuming and, as of now, clinically impractical.

Our study confirms the results of other studies that describe the adult preterm brain from a neuroradiological perspective [28–31]. We could confirm the smaller head size and the posterior dilatations of the lateral ventricles described by Aukland et al [28] and Bjuland et al [29], the lower volume of the posterior corpus callosum reported by Bjuland et al [29], the thinner corpus callosum and the white matter alterations described by Odberg et al [31]. Also, we could confirm previous reports of reduced deep grey matter volume in preterm born subjects [14, 29, 50, 51].

In contrast to some previous studies [14, 30, 50, 51] no decrease of white matter or cortical grey matter volumes could be found in our study. This is however not surprising as previous studies used finer volume measurements that highlight the regional distributions of white and grey matter. These studies reported both increases and decreases of regional volumes, which would cancel each other out in the coarse measures of whole grey or white matter volumes used in the present study. Studies that used similar coarse measures as reported here, such as for instance Bjuland et al [29], were also unable to find differences in normalized grey and white matter volumes between preterm and term born adults [29].

Our results confirm that in VP/VLBW the structural brain alterations related to preterm birth persist into adulthood. In our study, the highest diagnostic accuracy was achieved by adding the results of automatic and manual measures to the visual rating. The model using all quantitative and qualitative variables correctly classified 87% of cases. These results confirm the notion that the more parameters are considered in combination, the better the diagnostic accuracy.

The rapid development of computer-assisted methods of analysis and decision support through machine learning will augment and optimize human decision making and ultimately allow for precision medicine [52]. For instance, converting images to minable data and extraction of image features with deep learning algorithms - a field called radiomics - may ultimately generate predictive image-based phenotypes of disease and precision medicine. Our study demonstrates that the additional consideration of features like automatically determined volumes of lateral

ventricles, deep grey matter and posterior corpus callosum improves the accuracy of visual assessment.

Last but not least, the IQ of VP/VLBW group was lower than that of term born controls, which is in line with the results of Madzwamuse et al [25] who described the neuro-cognitive profile of the full BLS sample in adulthood [25]. Taken together, these results suggest that by adulthood, VP/VLBW as a group, do not outgrow their general cognitive deficits [25] nor, as our study confirmed, do they outgrow the structural brain alterations.

Limitations

One limitation of our study is the inclusion of a subset of the full cohort. However, as the main aim of the study was to investigate the preterm adult brain from a neuroradiological perspective, to avoid site effects we chose using all data that were collected in only one of the two study sites. Although the moderate sample size limits the statistical power to detect minor group differences, the aim of the present study was to evaluate the utility of these group differences from a neuroradiological point of view rather than to find group differences.

Further, our analysis included parameters most often affected in preterm born infants, children and adolescents, namely the brain with its ventricles, white matter and grey matter. However, future studies may include other parameters (biometric, cerebellar measurements, etc.) and use more complex machine learning algorithms to aid clinical decision-making.

Conclusion

Although prevalent, visual MR findings have low accuracy in diagnosing preterm birth related brain alterations in young adults. Automatic segmentation measurements (of lateral ventricles and deep grey matter), alone or in combination with manual measurements can improve diagnostic accuracy but they are time consuming and clinically impractical. It may be useful to ask for prematurity before initiating further diagnostics in subjects with thinner corpus callosum, dorsally deformed lateral wall of lateral ventricles, and “dirty” white matter.

References

1. Blencowe H, Cousens S, Oestergaard MZ, et al (2012) National, regional, and worldwide estimates of preterm birth rates in the year 2010 with time trends since 1990 for selected countries: a systematic analysis and implications. Lancet 379:2162–2172. https://doi.org/10.1016/S0140-6736(12)60820-4

2. Murphy M, McLoughlin G (2015) Born too soon: preterm birth in Europe trends, causes and prevention. Entre Nous 10–12

3. Zeitlin J, Szamotulska K, Drewniak N, et al (2013) Preterm birth time trends in Europe: a study of 19 countries. BJOG 120:1356–1365. https://doi.org/10.1111/1471-0528.12281

4. Martin JA, Hamilton BE, Ph D, et al (2017) Births: Final Data for 2015. Natl Vital Stat Rep 66:

5. Delnord M, Blondel B, Zeitlin J (2015) What contributes to disparities in the preterm birth rate in European countries? Curr Opin Obstet Gynecol 27:133–42. https://doi.org/10.1097/GCO.0000000000000156

6. Saigal S, Doyle LW (2008) An overview of mortality and sequelae of preterm birth from infancy to adulthood. Lancet 371:261–269. https://doi.org/10.1016/S0140-6736(08)60136-1

7. Allin M, Rooney M, Griffiths T, et al (2006) Neurological abnormalities in young adults born preterm. J Neurol Neurosurg Psychiatry 77:495–499. https://doi.org/10.1136/jnnp.2005.075465

8. Anderson PJ (2014) Neuropsychological outcomes of children born very preterm. Semin Fetal Neonatal Med 19:90–96. https://doi.org/10.1016/j.siny.2013.11.012

9. Baron IS, Rey-Casserly C (2010) Extremely preterm birth outcome: a review of four decades of cognitive research. Neuropsychol Rev 20:430–452. https://doi.org/10.1007/s11065-010-9132-z

10. Crump C, Sundquist K, Sundquist J, Winkleby MA (2011) Gestational age at birth and mortality in young adulthood. JAMA 306:1233–40. https://doi.org/10.1001/jama.2011.1331

11. Lawn JE, Blencowe H, Oza S, et al (2014) Every Newborn: progress, priorities, and potential beyond survival. Lancet Lond Engl 384:189–205. https://doi.org/10.1016/S0140-6736(14)60496-7

12. Hack M, Flannery DJ, Schluchter M, et al (2002) Outcomes in Young Adulthood for Very-Low-Birth-Weight Infants. N Engl J Med 346:149–157. https://doi.org/10.1056/NEJMoa010856

13. Rees S, Inder T (2005) Fetal and neonatal origins of altered brain development. Early Hum Dev 81:753–761. https://doi.org/10.1016/j.earlhumdev.2005.07.004

14. Bäuml JG, Daamen M, Meng C, et al (2014) Correspondence Between Aberrant Intrinsic Network Connectivity and Gray-Matter Volume in the Ventral Brain of Preterm Born Adults. Cereb Cortex. https://doi.org/10.1093/cercor/bhu133

15. Mathur AM, Neil JJ, Inder TE (2010) Understanding brain injury and neurodevelopmental disabilities in the preterm infant: the evolving role of advanced magnetic resonance imaging. Semin Perinatol 34:57–66. https://doi.org/10.1053/j.semperi.2009.10.006

16. Ment LR, Hirtz D, Hüppi PS (2009) Imaging biomarkers of outcome in the developing preterm brain. Lancet Neurol 8:1042–1055. https://doi.org/10.1016/S1474-4422(09)70257-1

17. Glass HC, Bonifacio SL, Peloquin S, et al (2010) Neurocritical care for neonates. Neurocrit Care 12:421–429. https://doi.org/10.1007/s12028-009-9324-7

18. Barkovich AJ, Raybaud C (2012) Pediatric Neuroimaging. Wolters Kluwer Health

19. Cheong JLY, Thompson DK, Wang HX, et al (2009) Abnormal white matter signal on MR imaging is related to abnormal tissue microstructure. AJNR Am J Neuroradiol 30:623–628. https://doi.org/10.3174/ajnr.A1399

20. Dubois J, Hu PS, Hüppi PS, Dubois J (2006) Diffusion tensor imaging of brain development. Semin Fetal Neonatal Med 11:489–497. https://doi.org/10.1016/j.siny.2006.07.006

21. Rutherford MA, Supramaniam V, Ederies A, et al (2010) Magnetic resonance imaging of white matter diseases of prematurity. Neuroradiology 52:505–521. https://doi.org/10.1007/s00234-010-0700-y

22. Degnan AJ, Wisnowski JL, Choi S, et al (2015) Alterations of resting state networks and structural connectivity in relation to the prefrontal and anterior cingulate cortices in late prematurity. Neuroreport 26:22–26. https://doi.org/10.1097/WNR.0000000000000296

23. Volpe JJ (2009) The encephalopathy of prematurity–brain injury and impaired brain development inextricably intertwined. Semin Pediatr Neurol 16:167–78. https://doi.org/10.1016/j.spen.2009.09.005

24. Basten M, Jaekel J, Johnson S, et al (2015) Preterm Birth and Adult Wealth: Mathematics Skills Count. Psychol Sci 26:1608–1619. https://doi.org/10.1177/0956797615596230

25. Eryigit Madzwamuse S, Baumann N, Jaekel J, et al (2014) Neuro-cognitive performance of very preterm or very low birth weight adults at 26 years. J Child Psychol Psychiatry n/a–n/a. https://doi.org/10.1111/jcpp.12358

26. Greene MF (2002) Outcomes of very low birth weight in young adults. N Engl J Med 346:146–8. https://doi.org/10.1056/NEJM200201173460302

27. Baumann N, Bartmann P, Wolke D (2016) Health-Related Quality of Life Into Adulthood After Very Preterm Birth. Pediatrics 137:. https://doi.org/10.1542/peds.2015-3148

28. Aukland SM, Elgen IB, Odberg MD, et al (2014) Ventricular dilatation in ex-prematures: only confined to the occipital region? MRI-based normative standards for 19-year-old ex-prematures without major handicaps. Acta Radiol 55:470–477. https://doi.org/10.1177/0284185113497476

29. Bjuland KJ, Rimol LM, Løhaugen GCC, Skranes J (2014) Brain volumes and cognitive function in very-low-birth-weight (VLBW) young adults. Eur J Paediatr Neurol 18:578–590. https://doi.org/10.1016/j.ejpn.2014.04.004

30. Nosarti C, Woo K, Walshe M, et al (2014) Preterm birth and structural brain alterations in early adulthood. NeuroImage Clin 6:180–191. https://doi.org/10.1016/j.nicl.2014.08.005

31. Odberg MD, Aukland SM, Rosendahl K, Elgen IB (2010) Cerebral MRI and cognition in nonhandicapped, low birth weight adults. Pediatr Neurol 43:258–62. https://doi.org/10.1016/j.pediatrneurol.2010.05.014

32. Breeman LD, Jaekel J, Baumann N, et al (2017) Neonatal predictors of cognitive ability in adults born very preterm: a prospective cohort study. Dev Med Child Neurol 59:477–483. https://doi.org/10.1111/dmcn.13380

33. Daamen M, Bäuml JG, Scheef L, et al (2015) Neural correlates of executive attention in adults born very preterm. NeuroImage Clin 9:581–591. https://doi.org/10.1016/j.nicl.2015.09.002

34. Jurcoane A, Daamen M, Scheef L, et al (2016) White matter alterations of the corticospinal tract in adults born very preterm and/or with very low birth weight. Hum Brain Mapp 37:289–299. https://doi.org/10.1002/hbm.23031

35. Riegel K, Ohrt B, Wolke D, Österlund K (1995) Die Entwicklung gefährdet geborener Kinder bis zum fünften Lebensjahr. (The development of at-risk children until the fifth year of life. The Arvo Ylppö longitudinal study in South Bavaria and South Finland). Stuttg Ferdinand Enke Verl

36. Wolke D, Meyer R (1999) Cognitive status, language attainment, and prereading skills of 6-year-old very preterm children and their peers: the Bavarian Longitudinal Study. Dev Med Child Neurol 41:94–109

37. Aster M von, Wechsler D (2009) Wechsler-Intelligenztest für Erwachsene : WIE ; Manual ; Übersetzung und Adaptation der WAIS-III von David Wechsler, 2., korr. Aufl. Pearson Assessment & Information, Frankfurt/M.

38. Tschampa HJ, Urbach H, Malter M, et al (2015) Magnetic resonance imaging of focal cortical dysplasia: Comparison of 3D and 2D fluid attenuated inversion recovery sequences at 3T. Epilepsy Res 116:8–14. https://doi.org/10.1016/j.eplepsyres.2015.07.004

39. Jenkinson M, Beckmann CF, Behrens TEJ, et al (2012) FSL. Neuroimage 62:782–790. https://doi.org/10.1016/j.neuroimage.2011.09.015

40. Ge Y, Grossman RI, Babb JS, et al (2003) Dirty-appearing white matter in multiple sclerosis: volumetric MR imaging and magnetization transfer ratio histogram analysis. AJNR Am J Neuroradiol 24:1935–1940

41. Rothman KJ, Greenland S, Lash TL (2008) Modern epidemiology, 3. ed. Wolters Kluwer, Lippincott Williams & Wilkins, Philadelphia, Pa.

42. Efron B, Tibshirani R (1993) An introduction to the bootstrap. Chapman & Hall, New York

43. Benjamini Y, Hochberg Y (1995) Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing. J R Stat Soc Ser B Methodol 57:289–300

44. Pandit AS, Ball G, Edwards a. D, Counsell SJ (2013) Diffusion magnetic resonance imaging in preterm brain injury. Neuroradiology 55:65–95. https://doi.org/10.1007/s00234-013-1242-x

45. Feldman HM, Lee ES, Loe IM, et al (2012) White matter microstructure on diffusion tensor imaging is associated with conventional magnetic resonance imaging findings and cognitive function in adolescents born preterm. Dev Med Child Neurol 54:809–814. https://doi.org/10.1111/j.1469-8749.2012.04378.x

46. Stewart AL, Rifkin L, Amess PN, et al (1999) Brain structure and neurocognitive and behavioural function in adolescents who were born very preterm. Lancet Lond Engl 353:1653–1657

47. Griffiths ST, Elgen IB, Chong WK, et al (2013) Cerebral Magnetic Resonance Imaging Findings in Children Born Extremely Preterm, Very Preterm, and at Term. Pediatr Neurol 49:113–118. https://doi.org/10.1016/j.pediatrneurol.2013.03.006

48. Ifflaender S, Rüdiger M, Konstantelos D, et al (2013) Prevalence of head deformities in preterm infants at term equivalent age. Early Hum Dev 89:1041–1047. https://doi.org/10.1016/j.earlhumdev.2013.08.011

49. Melbourne L, Murnick J, Chang T, et al (2015) Regional Brain Biometrics at Term-Equivalent Age and Developmental Outcome in Extremely Low-Birth-Weight Infants. Am J Perinatol 32:1177–1184. https://doi.org/10.1055/s-0035-1552936

50. Meng C, Bäuml JG, Daamen M, et al (2015) Extensive and interrelated subcortical white and gray matter alterations in preterm-born adults. Brain Struct Funct. https://doi.org/10.1007/s00429-015-1032-9

51. Nosarti C, Giouroukou E, Healy E, et al (2008) Grey and white matter distribution in very preterm adolescents mediates neurodevelopmental outcome. Brain 131:205–217. https://doi.org/10.1093/brain/awm282

52. Giger ML (2018) Machine Learning in Medical Imaging. J Am Coll Radiol. https://doi.org/10.1016/j.jacr.2017.12.028

Figure Legends

Figure 1. Analysis summary. On T1-w (A) and FLAIR images (C) we visually noted the presence of deformed lateral ventricles with billowed or bulged lateral walls of the posterior lateral ventricle (arrows in A) as well as the presence of white matter lesions or of regions with "dirty" appearing white matter (with intermediate signal intensity between lesions and normal-appearing white matter) (arrowheads in C). We also manually measured the widths of splenium (sw), isthmus (iw), body (cw) and genu (gw) of the corpus callosum (D) and automatically measured the volumes (B) of deep (red) and cortical grey matter (blue), white matter (yellow), lateral ventricles (green) as well as anterior and posterior corpus callosum (orange). Dotted vertical line in D delimits the anterior and posterior corpus callosum.

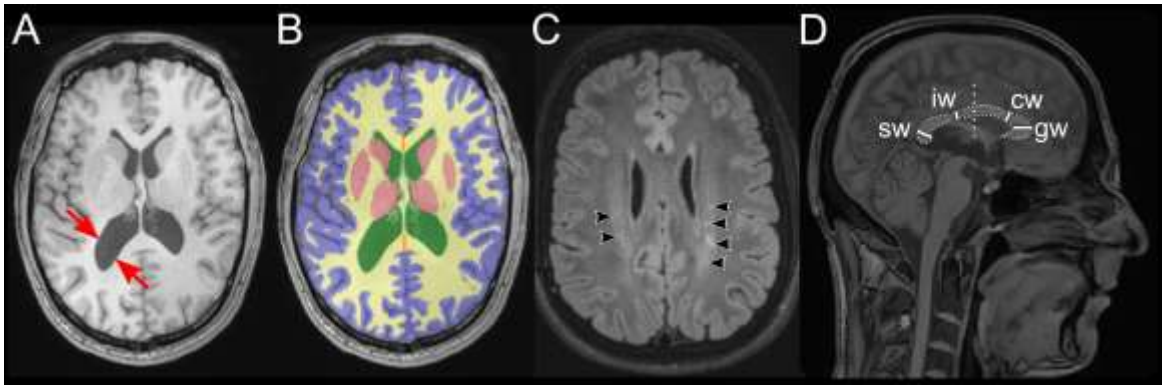

Figure 2. Results of the automatic segmentation (volumes), manual measurements (width) and visual rating analyses (aspect) including uncorrected P-values from t-tests or, for proportions, Fisher exact tests.

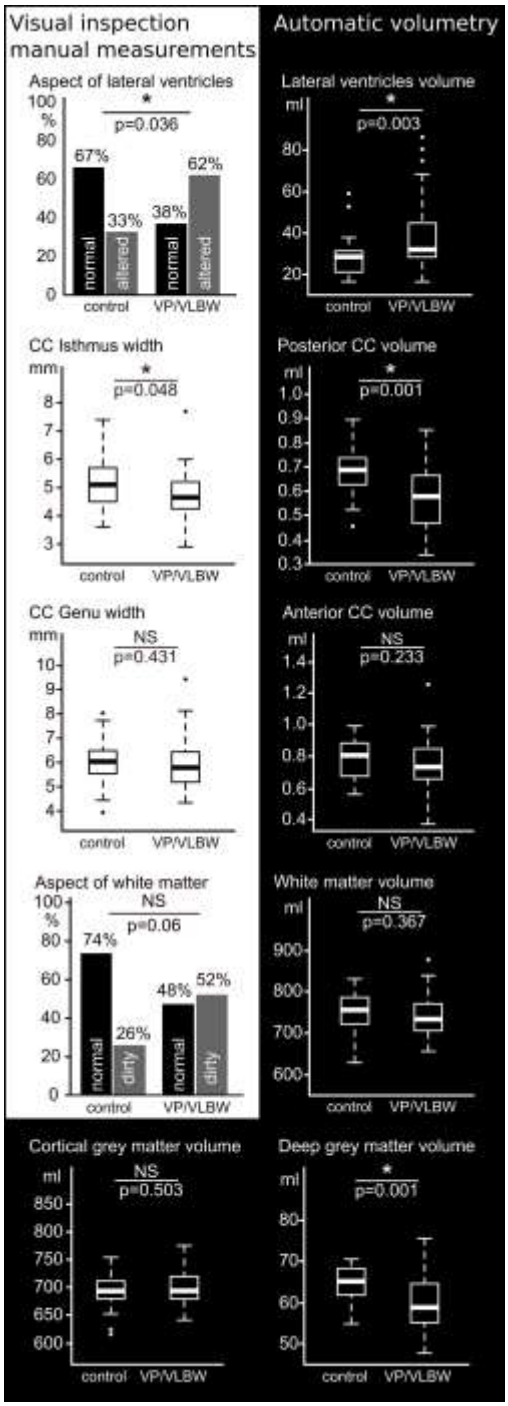

Figure 3. Exemplary images of lateral ventricles from all controls (on the right) and VP/VLBW patients (on the left). To enable easy visual evaluation, all images have been registered to a standard brain space, the Montreal Neurological Institute (MNI) space

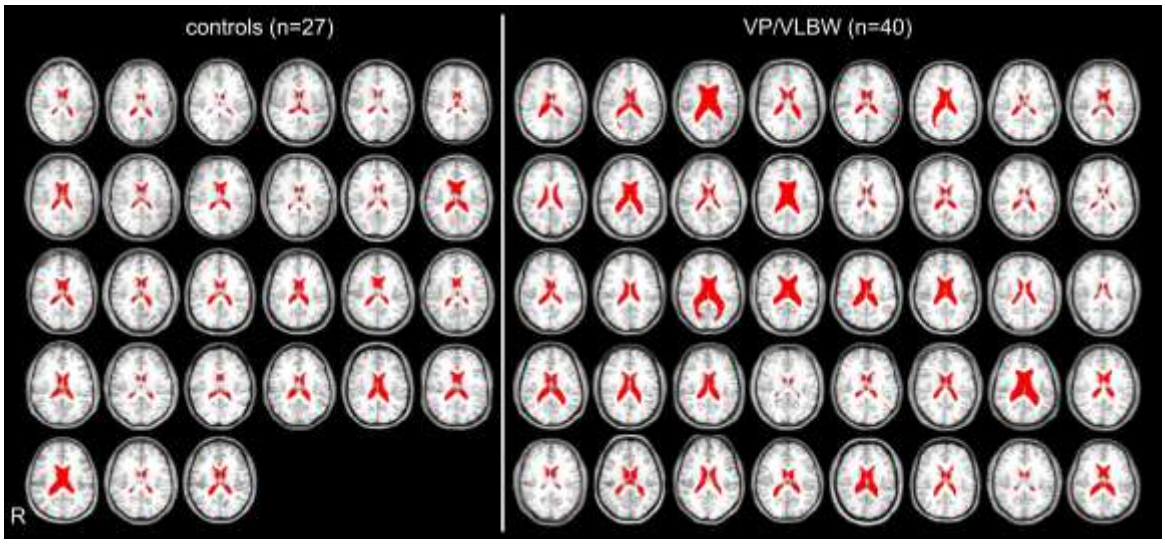

Table legends

Table 1. Background characteristics of VP/VLBW and term control study participants. P-values are derived from comparing control and preterm groups by t-tests or Fisher's exact test (for sex). Significant differences are bold face and marked with asterisk *

	control	**preterm**	**P-value control-preterm**
Adult age, mean(range), years	26.4 (25.4-27.7)	26.6 (25.7-27.7)	0.1
Adult Full-Scale IQ, mean(range)	107 (87.0-130)	97.3 (68.0-131)	0.002*
Sex, count(%)			0.4438
Female	8 (29.6%)	16 (40.0%)	
Male	19 (70.4%)	24 (60.0%)	
Birth weight, mean(range), g	3393 (2120-4200)	1351 (730-2070)	**<0.001***
Gestation age, mean(range), weeks	40.0 (38.0-42.0)	30.3 (25.0-36.0)	**<0.001***
Female	39.6 (38.0-41.0)	29.9 (25.0-36.0)	**<0.001***
Male	40.2 (39.0-42.0)	30.5 (27.0-35.0)	**<0.001***
P-value female-male	0.27	0.46	

Table 2. MRI findings. P-values are derived from comparing control and preterm groups by t-tests (continuous variables) and Fisher's exact tests (categorical variables) and are presented as values uncorrected and corrected for multiple comparisons with the false discovery rate method [43]. Results of manual measurements and visual evaluation are italicized, significant differences are bold face and marked with asterisk *.

	Controls (n=27)	VP/VLBW (n=40)	P uncorr	P corr
Lateral ventricles				
Volume (cm3)	27.6 (15.5-59.2)	38.0 (15.5-87.3)	**0.003***	**0.011***
Deformed shape	*9 (33.3%)*	*25 (62.5%)*	***0.036****	*0.109*
White matter				
Volume (cm3)	752 (629-831)	742 (656-877)	0.367	0.47
Altered ('dirty') aspect	*7 (25.9%)*	*21 (52.5%)*	*0.06*	*0.11*
Corpus callosum				
Volume (ant) (cm3)	0.79 (0.56-1.00)	0.75 (0.37-1.26)	0.233	0.35
Volume (post) (cm3)	0.68 (0.45-0.89)	0.56 (0.34-0.85)	**<0.001***	**0.001***
Genu width (mm)	*11.6 (8.20-14.2)*	*11.2 (7.10-16.3)*	*0.395*	*0.47*
Corpus ant width (mm)	*6.10 (3.90-8.00)*	*5.90 (4.30-9.40)*	*0.431*	*0.47*
Isthmus width (mm)	*5.19 (3.60-7.40)*	*4.73 (2.90-7.70)*	***0.048****	*0.112*
Splenium width (mm)	*12.3 (9.90-16.1)*	*11.6 (8.70-14.5)*	*0.087*	*0.149*
Grey matter				
Volume (deep nuclei) (cm3)	64.5 (54.9-70.7)	60.0 (47.7-75.7)	**0.001***	**0.003***
Volume (cortical) (cm3)	693 (615-754)	699 (640-775)	0.503	0.503

Table 3. Diagnostic performance of models including MRI variables; optimal cut-point was set to 0.5 to maximize simultaneously both sensitivity and specificity. P-Value represents the difference from the full model, with all MRI variables. Data are percent parameter scores (and 95% confidence intervals in parentheses). Bootstrapped AUC is the average of the estimated AUC across 10'000 simulations. Abbreviations: PPV positive predictive value, NPV negative predictive value, ROC receiver operating characteristic, AUC area under the curve.

[1] (volume of deep grey matter, ventricular volume and white matter aspect)

Model / Performance	Full model (all MRI variables)	Automatic segmentation	Quantification (automatic and manual)	Selected variables[1], stepwise	Manual and visual evaluation	Manual measurements	Visual inspection
Sensitivity	88%(73-96%)	88%(73-96%)	85%(70-94%)	85%(70-94%)	85%(70-94%)	82%(67-93%)	78%(62-89%)
Specificity	85%(66-96%)	78%(58-91%)	74%(54-89%)	74%(54-89%)	56%(35-75%)	41%(22-61%)	41%(22-61%)
PPV	90%(76-97%)	85%(71-94%)	83%(68-93%)	83%(68-93%)	74%(59-86%)	67%(52-80%)	66%(51-79%)
NPV	82%(63-94%)	81%(61-93%)	77%(56-91%)	77%(56-91%)	71%(48-89%)	61%(36-83%)	55%(32-77%)
Correctly classified (accuracy)	87%(76-94%)	84%(73-92%)	81%(69-89%)	81%(69-89%)	73%(61-83%)	66%(53-77%)	63%(50-74%)
ROC AUC	0.95	0.88	0.92	0.81	0.76	0.67	0.69
P-value	-	0.32	0.97	0.02	0.008	0.0003	0.0002
Logistic Regression: Internally-Validated (10000 bootstrap samples)							
AUC 'in-sample'	**0.98**	**0.91**	**0.96**	**0.85**	**0.79**	**0.71**	**0.7**

Supplementary Table 1. MR imaging parameters. Abbreviations: FA flip angle; IR inversion recovery; SENSE Sensibility Encoding (parallel acquisition)

	T1-w	**FLAIR**	**T2-w**
Sequence	3D Fast Field Echo (with IR prepulse)	3D Fluid attenuated IR	3D Fast Spin Echo
TR (ms)	7.7	4800	2500
TE(ms)	3.93	320	363
TI(ms)	1300	1600	-
FA	15°	90°	90°
Slices (partitions)	180	360	360
SENSE factor	2	5	4
matrix	256 x 256	512x512	512x512
FOV	256 x 256	256 x 256	256 x 256
Spectral fat suppression	-	yes	yes
reconstructed voxel size	1 mm^3	0.5 mm^3	0.5 mm^3

Supplementary Table 2. Details of all the logistic regression models analyzed in the manuscript including the variables included in the models, the odds ratio (OR) and the P-Value (P). The interpretation of the output in the metric of odds ratio, is as follows: "For a unit change in a variable, the odd for an event is expected to change by a factor of OR, holding all other variables constant." For instance, for the full model: the odd of preterm birth increases by a factor of 1.15 (or 15%) per cm^3 increase of volume of lateral ventricles (p=0.076), or, the odd of preterm birth decreases by 0.737 (or 26.3%) per cm^3 increase of deep grey matter nuclei (p=0.092).

Model / Variable name	**all MRI variables (the full model)**		**automatic segmentation**		**quantification (automatic and manual)**		**selected variables[1], stepwise**		**manual and visual evaluation**		**manual measurements**		**visual inspection**	
	OR	P	OR	P	OR	P	OR	P	OR	P	OR	P	OR	P
Volume (cm3) of Lateral Ventricles	1.151	0.076	1.09	0.037	1.128	0.039	1.082	0.02	-	-	-	-	-	-
Volume (cm3) of cortical grey matter	1.057	0.03	1.034	0.038	1.054	0.014	0.878	0.026	-	-	-	-	-	-
Volume (cm3) of White Matter	1.035	0.835	0.979	0.868	1.035	0.817	-	-	-	-	-	-	-	-
Volume (cm3) of deep grey matter nuclei	0.737	0.092	0.786	0.03	0.685	0.029	-	-	-	-	-	-	-	-
Volume (mm3) of anterior Corpus Callosum	1.049	0.21	1.041	0.169	1.04	0.216	-	-	-	-	-	-	-	-
Volume (mm3) of posterior Corpus Callosum	1.028	0.446	1.028	0.342	1.023	0.47	-	-	-	-	-	-	-	-
Volume (mm3) of whole Corpus Callosum	0.958	0.229	0.965	0.205	0.967	0.265	-	-	-	-	-	-	-	-
Width (mm) of Genu of the Corpus Callosum	1.458	0.341	-	-	1.173	0.628	-	-	1.18	0.392	1.014	0.933	-	-
Width (mm) of anterior Corpus Callosum	2.627	0.193	-	-	2.135	0.172	-	-	1.143	0.75	1.376	0.402	-	-
Width (mm) of posterior Corpus Callosum (Isthmus)	0.206	0.058	-	-	0.209	0.033	-	-	0.488	0.101	0.469	0.069	-	-
Width (mm) of Splenium of the Corpus Callosum	1.735	0.12	-	-	1.325	0.338	-	-	0.789	0.225	0.761	0.13	-	-
Deformed lateral ventricles	2.847	0.422	-	-	-	-	-	-	3.049	0.065	-	-	2.775	0.056
Altered ('dirty') aspect of White Matter	7.153	0.058	-	-	-	-	4.432	0.026	2.674	0.09	-	-	3.096	0.042